Investigating Urban Cycling Demand Utilising Crowdsourced Data

V. Petraki*, A. Kanakis, E. Koliou, G. Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens
5 Heroon Polytechniou str., 15773, Athens, Greece

Keywords: cycling, crowdsourced data, STRAVA, spatial analysis, lognormal regression.

1 INTRODUCTION

Efforts are being made globally to promote active travel modes, and reduce reliance on motorized traffic in a broader effort to address congestion and air pollution, reduce energy dependency, and enhance equity and accessibility for all (Pucher & Buehler, 2021). To plan for active modes and assess policies to promote cycling, city and transport authorities rely on data. However, for cycling, data is hardly available, especially in comparison to other modes, and usually comes from several detectors/counters covering a small period and a limited area. Exploiting novel data sources, such as crowdsourced data, addresses data gaps (Nelson et al., 2021). This research focuses on "Strava Metro" cycling trips data and aims to demonstrate a framework to utilize these data to understand and model spatiotemporal aspects of cycling trips. The objective is to develop a mathematical model using open data to identify key factors influencing cycling demand, identifying environments that support both higher demand and should be prioritized by policymakers and transport engineers for targeted cycling safety improvements.

2 METHOD

The study collected cycling activity data from "Strava Metro" platform recorded for the years 2021 and 2022, on 11,446 road segments of two municipalities in Athens, Greece, namely Chalandri and Vrillisia (see Figure 1). These two adjacent municipalities were selected for their developed cycling network, and the respective higher cycling traffic. This data was spatially represented using a Geographic Information System and enriched with additional road characteristics (e.g., road types, segment lengths) and surrounding land uses sourced from OpenStreetMap and processed with QGIS.

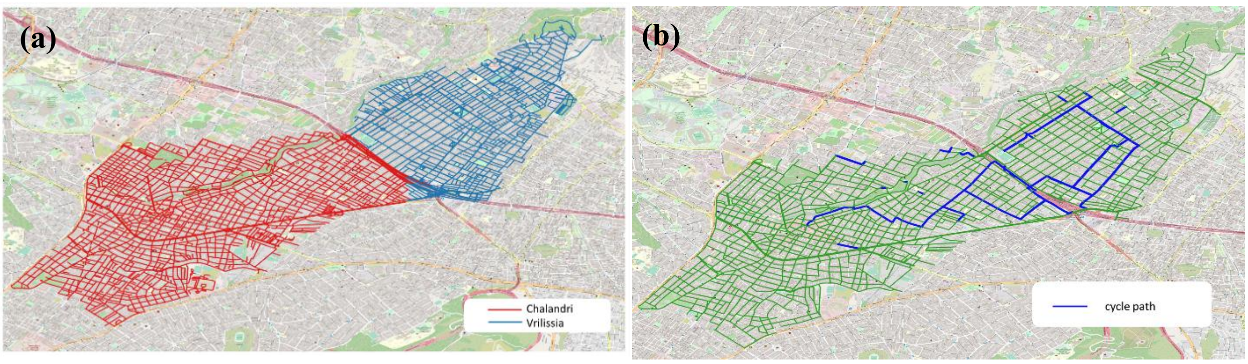

Figure 1: (a) Study Areas (b) Cycle path on study areas

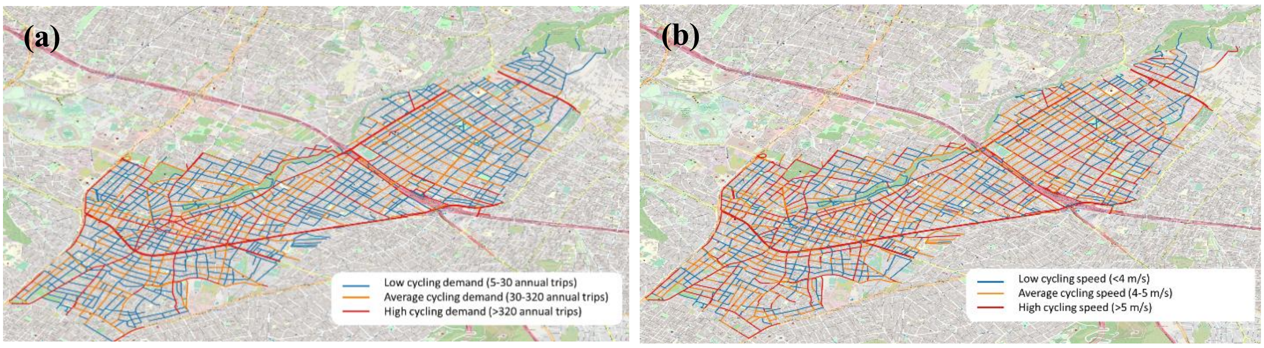

Figure 2: (a) Cycling demand heatmap (b) Cycling speed heatmap

3 RESULTS

he following table presents the general log-normal regression model for both areas, achieving a satisfactory $R^2 = 0.66$. Also, all independent variables are statistically significant at a 95% confidence level. To complement the developed models, elasticity analyses were conducted as well.

Table 1: Results of lognormal cycling demand model.

Variable		Reference Category	Estimate	t. value	elasticity	
					e	e*
Intercept		-	5.8204	55.513	-	-
Average Cycling Speed		-	0.0510	7.509	0.002	2.656
Road Category	pedestrian zone	cycleway	-2.3461	19.675	-0.151	-25.555
	residential		-2.2292	29.326	-0.045	-7.550
	motorway		-1.9854	4.254	-0.119	-20.169
	primary		2.7203	8.730	0.024	4.067
	secondary		2.4350	26.258	0.024	4.087
	tertiary		0.2683	3.180	0.006	1.000
	unclassified		-2.8992	11.166	-0.237	-40.083
Road Segment Length (m)		-	-0.0019	3.822	-0.001	-1.198
Proximity to land uses (m)	proximity to bicycle parking spots	-	-0.0004	8.898	-0.004	-4.490
	proximity to education areas	-	0.0004	2.289	0.001	1.000
	proximity to leisure parks	-	0.0008	4.747	0.001	1.442
	proximity to public transport	-	-0.0008	3.213	-0.001	-1.087
2022 year		2021 year	-0.7537	55.070	-0.019	-3.195

4 DISCUSSION

The findings are strongly aligned with international literature, confirming that infrastructure quality, connectivity, and proximity to cycling-supportive amenities significantly influence cycling behavior (Yan et al., 2024; Fosgerau et al., 2023).

The analysis reveals that higher average cycling speeds on road segments are associated significantly with increased cycling activity. A 1% increase in the recorded cycling speed resulted in a 2.7% increase in cycling trips. However, high cycling speed, is a key indicator of cycling risk. This suggests that segments with increased cycling trips may also be those where safety interventions are most urgently needed.

Interestingly, the analysis revealed that primary, secondary, and tertiary roads are associated with higher cycling demand compared to cycleways. This suggests that many cyclists choose these roads, possibly because they offer more direct or continuous routes through the road network. However, these roads are typically shared with motorized traffic and lack dedicated cycling infrastructure, raising serious safety concerns. Segment length had a negative effect; longer segments may deter cycling due to increased exposure to traffic and a lack of intermediate safe crossing points.

Furthermore, the analysis found lower cycling activity in 2022 compared to 2021. This corresponds with COVID-19 restrictions, which initially drove a surge in outdoor activities, while the return to pre-pandemic travel habits led to a reduction (Kraus & Koch, 2021).

Finally, accessibility to supportive amenities emerged as another crucial determinant. Cycling trips were more frequent when routes were closer to bicycle parking facilities and public transport stops. Proximity to these amenities offers cyclists greater convenience fostering a higher likelihood of bicycle use, but they also require careful design to ensure safe access.

5 CONCLUSION

This study developed a model for investigating cycling activity in two Athens municipalities using high-resolution crowdsourced data. The findings highlight how infrastructure type, cycling speed, and proximity to urban amenities affect cycling. Cyclists are more likely to use major roads over cycleways, revealing a potential mismatch between demand and safe infrastructure. The positive association between cycling demand and average speed, commonly used as a risk factor, suggests that high-demand segments may also pose safety concerns. Proximity to amenities like bike parking, schools, and parks is positively associated with cycling activity, emphasizing the need to connect infrastructure with everyday destinations. These insights show how cycling demand data can guide infrastructure and safety priorities. Future research could apply this method to other cities, include seasonal or real-time data, and combine demand modeling with crash records to support safer cycling networks.

6 REFERENCES

Buehler, R., & Pucher, J. (Eds.). (2021). Cycling for sustainable cities. MIT Press.

Nelson, T., Ferster, C., Laberee, K., Fuller, D., & Winters, M. (2021). Crowdsourced data for bicycling research and practice. Transport reviews, 41(1), 97-114.

Yan, H., Maat, K., & van Wee, B. (2024). Cycling speed variation: a multilevel model of characteristics of cyclists, trips and route tracking points. Transportation, 1-30.

Fosgerau, M., Łukawska, M., Paulsen, M., & Rasmussen, T. K. (2023). Bikeability and the induced demand for cycling. Proceedings of the National Academy of Sciences, 120(16), e2220515120.

Kraus, S., & Koch, N. (2021). Provisional COVID-19 infrastructure induces large, rapid increases in cycling. Proceedings of the National Academy of Sciences, 118(15), e2024399118.